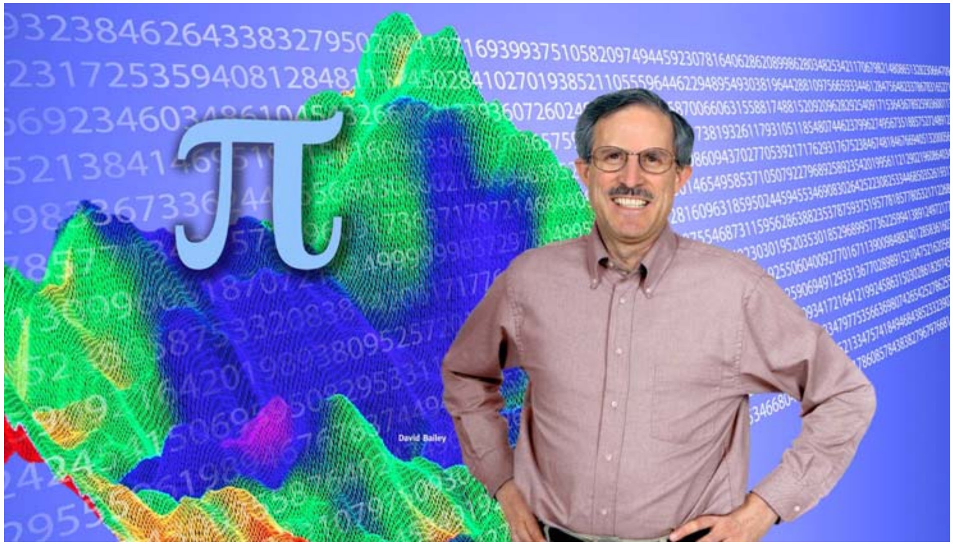
Experimental Mathematics and the Normality of Pi

David H Bailey

Lawrence Berkeley National Lab



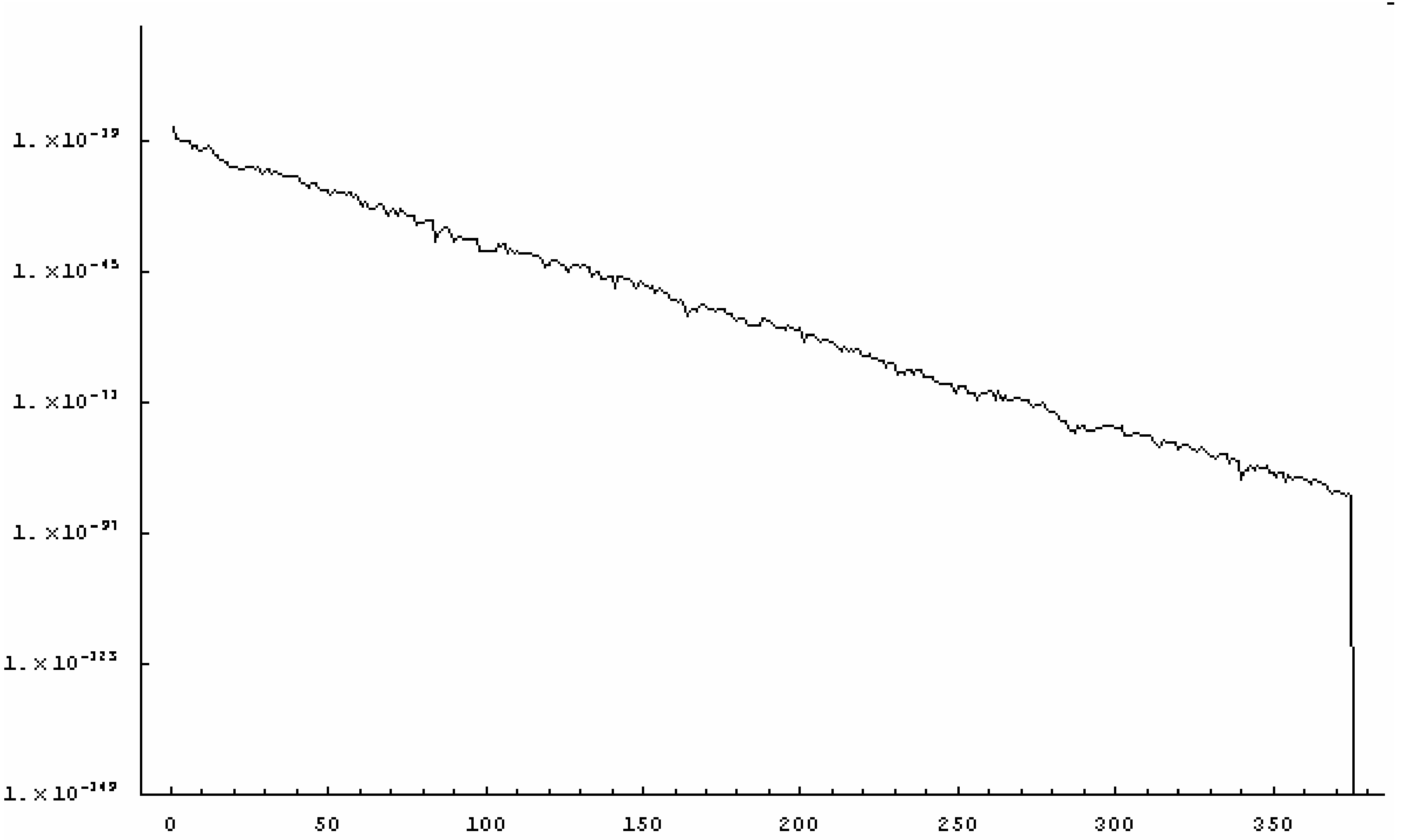
The PSLQ Integer Relation Algorithm

Let (x_n) be a vector of real numbers. An integer relation algorithm finds integers (a_n) such that

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0$$

- At the present time, the PSLQ algorithm of mathematician-sculptor Helaman Ferguson is the best-known integer relation algorithm.
- PSLQ was named one of ten “algorithms of the century” by *Computing in Science and Engineering*.
- High precision arithmetic software is required: at least d x n digits, where d is the size (in digits) of the largest of the integers a_k.

Decrease of min $|x_i|$ in PSLQ



Some Supercomputer-Class PSLQ Solutions

- Identification of B_4, the fourth bifurcation point of the logistic iteration.
  - Integer relation of size 121; 10,000 digit arithmetic.
- Identification of Apery sums.
  - 15 integer relation problems, with size up to 118, requiring up to 5,000 digit arithmetic.
- Identification of Euler-zeta sums.
  - Hundreds of integer relation problems, each of size 145 and requiring 5,000 digit arithmetic.
  - Run on IBM SP parallel system.
- Finding relation involving root of Lehmer's polynomial.
  - Integer relation of size 125; 50,000 digit arithmetic.
  - Utilizes 3-level, multi-pair parallel PSLQ program.
  - Run on IBM SP using ARPEC; 16 hours on 64 CPUs.

LBNL's Arbitrary Precision Computation (ARPREC) Package

- Low-level routines written in C++.
- C++ and F-90 translation modules permit use with existing programs with only minor code changes.
- Double-double (32 digits), quad-double, (64 digits) and arbitrary precision (>64 digits) available.
- Special routines for extra-high precision (>1000 dig).
- Includes common math functions: sqrt, cos, exp, etc.
- PSLQ, root finding, numerical integration.
- An interactive "Experimental Mathematician's Toolkit" employing this software is also available.

Available at: **http://www.experimentalmath.info**

The Tanh-Sinh Numerical Integration Scheme

Given f(x) defined on (-1,1), substitute x = g(t), where g(t) = tanh (π/2 * sinh t). Then we can write

$$\int_{-1}^{1} f(x)\,dx \;=\; \int_{-\infty}^{\infty} f(g(t))g'(t)\,dt \;\approx\; h\sum_{-N}^{N} w_j f(x_j)$$

Here x_j = g(hj) and w_j = g'(hj).

As a consequence of the Euler-Maclaurin summation formula, such approximations converge very rapidly – for most functions, reducing h by half produces twice as many correct digits.

The tanh-sinh scheme works well even for functions with infinite derivatives or blow-up singularities at endpoints.

Test Integrals

$$1: \int_0^1 t\log(1+t)\,dt = 1/4$$

$$2: \int_0^1 t^2 \arctan t\,dt = (\pi - 2 + 2\log 2)/12$$

$$3: \int_0^{\pi/2} e^t \cos t\,dt = (e^{\pi/2} - 1)/2$$

$$4: \int_0^1 \frac{\arctan(\sqrt{2+t^2})}{(1+t^2)\sqrt{2+t^2}}\,dt = 5\pi^2/96$$

$$5: \int_0^1 \sqrt{t}\log t\,dt = -4/9$$

$$6: \int_0^1 \sqrt{1-t^2}\,dt = \pi/4$$

$$7: \int_0^1 \frac{t}{\sqrt{1-t^2}}dt = 1$$

$$8: \int_0^1 \log t^2\,dt = 2$$

$$9: \int_0^{\pi/2} \log(\cos t)\,dt = -\pi\log(2)/2$$

$$10: \int_0^{\pi/2} \sqrt{\tan t}\,dt = \pi\sqrt{2}/2$$

$$11: \int_0^\infty \frac{1}{1+t^2}\,dt = \pi/2$$

$$12: \int_0^\infty \frac{e^{-t}}{\sqrt{t}}\,dt = \sqrt{\pi}$$

$$13: \int_0^\infty e^{-t^2/2}\,dt = \sqrt{\pi/2}$$

$$14: \int_0^\infty e^{-t}\cos t\,dt = 1/2$$

Quadratic Convergence with Tanh-Sinh Quadrature

BERKELEY LAB

Level	Prob. 1	Prob. 2	Prob. 3	Prob. 4	Prob. 5	Prob. 6	Prob. 7
1	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-5}	10^{-5}	10^{-6}
2	10^{-11}	10^{-11}	10^{-9}	10^{-9}	10^{-12}	10^{-12}	10^{-12}
3	10^{-24}	10^{-19}	10^{-21}	10^{-18}	10^{-28}	10^{-25}	10^{-26}
4	10^{-51}	10^{-38}	10^{-49}	10^{-36}	10^{-62}	10^{-50}	10^{-49}
5	10^{-98}	10^{-74}	10^{-106}	10^{-73}	10^{-129}	10^{-99}	10^{-98}
6	10^{-195}	10^{-147}	10^{-225}	10^{-145}	10^{-265}	10^{-196}	10^{-194}
7	10^{-390}	10^{-293}	10^{-471}	10^{-290}	10^{-539}	10^{-391}	10^{-388}
8	10^{-777}	10^{-584}	10^{-974}	10^{-582}		10^{-779}	10^{-777}

Level	Prob. 8	Prob. 9	Prob. 10	Prob. 11	Prob. 12	Prob. 13	Prob. 14
1	10^{-5}	10^{-4}	10^{-6}	10^{-2}	10^{-2}	10^{-1}	10^{-1}
2	10^{-12}	10^{-11}	10^{-12}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
3	10^{-29}	10^{-24}	10^{-25}	10^{-11}	10^{-9}	10^{-6}	10^{-5}
4	10^{-62}	10^{-50}	10^{-48}	10^{-22}	10^{-15}	10^{-9}	10^{-8}
5	10^{-130}	10^{-97}	10^{-98}	10^{-45}	10^{-28}	10^{-19}	10^{-14}
6	10^{-266}	10^{-195}	10^{-194}	10^{-91}	10^{-50}	10^{-37}	10^{-26}
7	10^{-540}	10^{-389}	10^{-388}	10^{-182}	10^{-92}	10^{-66}	10^{-48}
8		10^{-777}	10^{-777}	10^{-365}	10^{-170}	10^{-126}	10^{-88}
9				10^{-731}	10^{-315}	10^{-240}	10^{-164}
10					10^{-584}	10^{-457}	10^{-304}
11						10^{-870}	10^{-564}

At level k, $h = 2^{-k}$. I.e., each level halves h and doubles N, the # of abscissas.

Experimental Result Using PSLQ and Tanh-Sinh Quadrature - Example 1

Let

$$C(a) \;=\; \int_0^1 \frac{\arctan\sqrt{x^2+a^2}}{(x^2+1)\sqrt{x^2+a^2}}dx$$

Then PSLQ yields

$$\begin{aligned}
C(0) &= (\pi \log 2)/8 + G/2 \\
C(1) &= \pi/4 - \pi\sqrt{2}/2 + 3\sqrt{2}/2 \cdot \arctan\sqrt{2} \\
C(\sqrt{2}) &= 5\pi^2/96
\end{aligned}$$

Several general results have now been proven, including

$$\int_0^\infty \frac{\arctan\sqrt{x^2+a^2}\,dx}{(x^2+1)\sqrt{x^2+a^2}} \;=\; \frac{\pi}{2\sqrt{a^2-1}}\left(2\arctan\sqrt{a^2-1} - \arctan\sqrt{a^4-1}\right)$$

Example 2

$$\frac{2}{\sqrt{3}}\int_0^1 \frac{\log^6 x \arctan[x\sqrt{3}/(x-2)]}{x+1}\,dx \;=\;$$

$$\frac{1}{81648}\Big(-229635 L_{-3}(8) + 29852550 L_{-3}(7)\log 3$$

$$-1632960 L_{-3}(6)\pi^2 + 27760320 L_{-3}(5)\zeta(3)$$

$$-275184 L_{-3}(4)\pi^4 + 36288000 L_{-3}(3)\zeta(5)$$

$$-30008 L_{-3}(2)\pi^6 - 57030120 L_{-3}(1)\zeta(7)\Big)$$

where

$$L_{-3}(s) \;=\; \sum_{n=1}^{\infty}\left[1/(3n-2)^s - 1/(3n-1)^s\right]$$

is the Dirichlet series.

Example 3

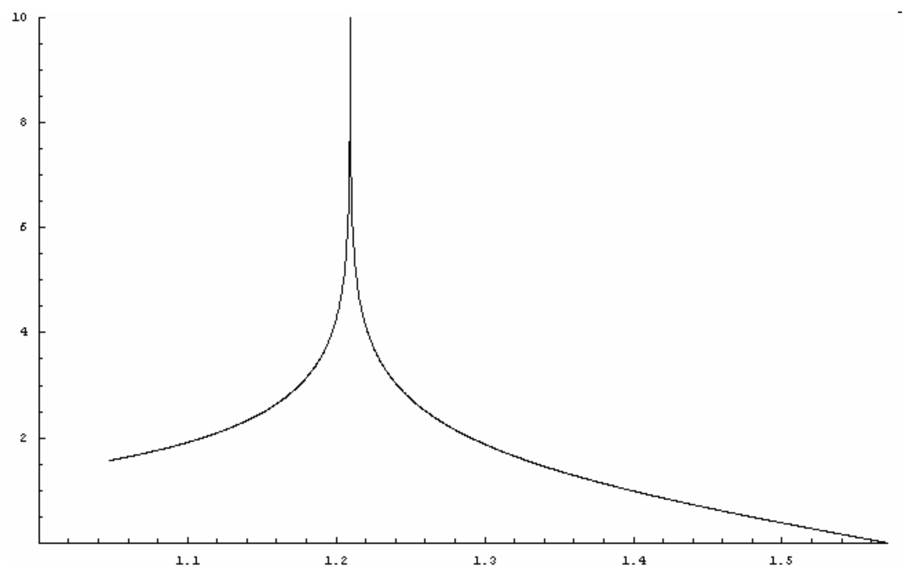
$$\frac{24}{7\sqrt{7}} \int_{\pi/3}^{\pi/2} \log \left| \frac{\tan t + \sqrt{7}}{\tan t - \sqrt{7}} \right| dt$$

$$\stackrel{?}{=} \sum_{n=0}^{\infty} \left[\frac{1}{(7n+1)^2} + \frac{1}{(7n+2)^2} - \frac{1}{(7n+3)^2} + \frac{1}{(7n+4)^2} - \frac{1}{(7n+5)^2} - \frac{1}{(7n+6)^2} \right]$$

This arises in mathematical physics, from analysis of the volumes of ideal tetrahedra in hyperbolic space.

This “identity” has now been verified numerically to 20,000 digits, but no proof is known.

Note that the integrand function has a nasty singularity.



Example 4

Define

$$J_n = \int_{n\pi/60}^{(n+1)\pi/60} \log \left| \frac{\tan t + \sqrt{7}}{\tan t - \sqrt{7}} \right| dt$$

Then

$$\begin{aligned} 0 \stackrel{?}{=} & -2J_2 - 2J_3 - 2J_4 + 2J_{10} + 2J_{11} + 3J_{12} \\ & +3J_{13} + J_{14} - J_{15} - J_{16} - J_{17} - J_{18} \\ & -J_{19} + J_{20} + J_{21} - J_{22} - J_{23} + 2J_{25} \end{aligned}$$

This has been verified to over 1000 digits. The interval in J_{23} includes the singularity.

Example 5 (Jan 2006)

The following integrals are related to the Ising theory of mathematical physics:

$$C_n = \frac{4}{n!} \int_0^\infty \cdots \int_0^\infty \frac{1}{\left(\sum_{j=1}^n (u_j + 1/u_j)\right)^2} \frac{du_1}{u_1} \cdots \frac{du_n}{u_n}$$

We first showed that this can be transformed to a 1-D integral:

$$C_n = \frac{2^n}{n!} \int_0^\infty t K_0^n(t) \; dt$$

where K_0 is a modified Bessel function. We then computed 500-digit numerical values, from which we found these results (now proven):

$$C_3 = L_{-3}(2) = \sum_{n \geq 0} \left(\frac{1}{(3n+1)^2} - \frac{1}{(3n+2)^2} \right)$$

$$C_4 = 14\zeta(3)$$

$$\lim_{n \to \infty} C_n = 2e^{-2\gamma}$$

Cautionary Example

These constants agree to 42 decimal digit accuracy, but are NOT equal:

$$\int_0^\infty \cos(2x) \prod_{n=0}^{\infty} \cos(x/n)\, dx \quad =$$

$$0.392699081698724154807830422909937860524645434187231\ldots$$

$$\frac{\pi}{8} \quad =$$

$$0.392699081698724154807830422909937860524646174921888\ldots$$

Computing this integral is nontrivial, due to difficulty in evaluating the integrand function to high precision.

Fascination With Pi

Newton (1670):

"I am ashamed to tell you to how many figures I carried these computations, having no other business at the time."

Carl Sagan (1986):

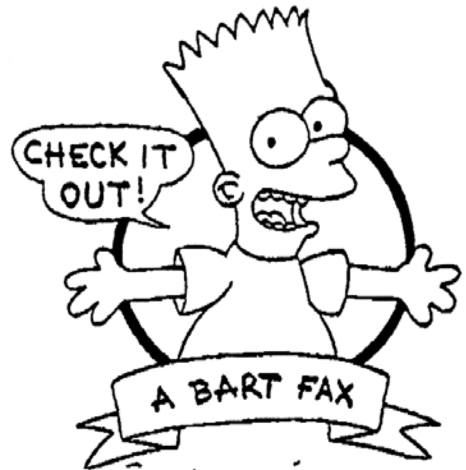
In his book "Contact," the lead scientist (played by Jodie Foster in the movie) looked for patterns in the digits of pi.

Carl E. Sagan

Wall Street Journal (Mar 15, 2005):

"Yesterday was Pi Day: March 14, the third month, 14th day. As in 3.14, roughly the ratio of a circle's circumference to its diameter…"

Fax from "The Simpsons" Show



TO: David Bailey

FROM: Jacqueline Atkins

DATE: 10/9/92

NUMBER OF PAGES: 1

FAX (310) 203-3852

PHONE (310) 203-3959

A Professor at UCLA told me that you might be able to give me the answer to: What is the 40,000th digit of Pi?

We would like to use the answer in our show. Can you help?

Peter Borwein's Observation

In 1996, Peter Borwein of SFU in Canada observed that the following well-known formula for $\log_e 2$

$$\log 2 = \sum_{n=1}^{\infty} \frac{1}{n2^n} = 0.693147180559945309 42\ldots$$

leads to a simple scheme for computing binary digits at an arbitrary starting position (here {} denotes fractional part):

$$\begin{aligned} \left\{2^d \log 2\right\} &= \left\{\sum_{n=1}^{d} \frac{2^{d-n}}{n}\right\} + \sum_{n=d+1}^{\infty} \frac{2^{d-n}}{n} \\ &= \left\{\sum_{n=1}^{d} \frac{2^{d-n} \bmod n}{n}\right\} + \sum_{n=d+1}^{\infty} \frac{2^{d-n}}{n} \end{aligned}$$

Fast Exponentiation Mod n

The exponentiation (2^{d-n} mod n) in this formula can be evaluated very rapidly by means of the binary algorithm for exponentiation, performed modulo n:

Example:

$3^{17} = (((3^2)^2)^2)^2 \times 3 = 129140163$

In a similar way, we can evaluate:

$3^{17} \bmod 10 = ((((3^2 \bmod 10)^2 \bmod 10)^2 \bmod 10)^2 \bmod 10) \times 3 \bmod 10$

$3^2 \bmod 10 = 9$

$9^2 \bmod 10 = 1$

$1^2 \bmod 10 = 1$

$1^2 \bmod 10 = 1$

$1 \times 3 = 3$ Thus $3^{17} \bmod 10 = 3$.

Note: we never have to deal with integers larger than 81.

The BBP Formula for Pi

In 1996, Simon Plouffe, using DHB's PSLQ program, discovered this formula for pi:

$$\pi = \sum_{k=0}^{\infty} \frac{1}{16^k} \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right)$$

Indeed, this formula permits one to directly calculate binary or hexadecimal (base-16) digits of π beginning at an arbitrary starting position n, without needing to calculate any of the first n-1 digits.

Proof of the BBP Formula

$$\int_0^{1/\sqrt{2}} \frac{x^{j-1}\,dx}{1-x^8} = \int_0^{1/\sqrt{2}} \sum_{k=0}^{\infty} x^{8k+j-1}\,dx = \frac{1}{2^{j/2}} \sum_{k=0}^{\infty} \frac{1}{16^k(8k+j)}$$

Thus

$$\begin{aligned}
\sum_{k=0}^{\infty} \frac{1}{16^k} \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right) \\
&= \int_0^{1/\sqrt{2}} \frac{(4\sqrt{2} - 8x^3 - 4\sqrt{2}x^4 - 8x^5)\,dx}{1-x^8} \\
&= \int_0^1 \frac{16(4 - 2y^3 - y^4 - y^5)\,dy}{16-y^8} \\
&= \int_0^1 \frac{16(y-1)\,dy}{(y^2-2)(y^2-2y+2)} \\
&= \int_0^1 \frac{4y\,dy}{y^2-2)} - \int_0^1 \frac{(4y-8)\,dy}{y^2-2y+2} \\
&= \pi
\end{aligned}$$

Calculations Using the BBP Algorithm

Position	Hex Digits of Pi Starting at Position	
10^6	26C65E52CB4593	
10^7	17AF5863EFED8D	
10^8	ECB840E21926EC	
10^9	85895585A0428B	
10^{10}	921C73C6838FB2	
10^{11}	9C381872D27596	
1.25×10^{12}	07E45733CC790B	[1]
2.5×10^{14}	E6216B069CB6C1	[2]

[1] Babrice Bellard, France, 1999

[2] Colin Percival, Canada, 2000

Some Other Similar New Identities

$$\pi\sqrt{3} = \frac{9}{32}\sum_{k=0}^{\infty}\frac{1}{64^k}\left(\frac{16}{6k+1}-\frac{8}{6k+2}-\frac{2}{6k+4}-\frac{1}{6k+5}\right)$$

$$\pi^2 = \frac{1}{8}\sum_{k=0}^{\infty}\frac{1}{64^k}\left(\frac{144}{(6k+1)^2}-\frac{216}{(6k+2)^2}-\frac{72}{(6k+3)^2}-\frac{54}{(6k+4)^2}+\frac{9}{(6k+5)^2}\right)$$

$$\pi^2 = \frac{2}{27}\sum_{k=0}^{\infty}\frac{1}{729^k}\left(\frac{243}{(12k+1)^2}-\frac{405}{(12k+2)^2}-\frac{81}{(12k+4)^2}-\frac{27}{(12k+5)^2}\right.$$
$$\left.-\frac{72}{(12k+6)^2}-\frac{9}{(12k+7)^2}-\frac{9}{(12k+8)^2}-\frac{5}{(12k+10)^2}+\frac{1}{(12k+11)^2}\right)$$

$$6\sqrt{3}\,\mathrm{arctan}\left(\frac{\sqrt{3}}{7}\right) = \sum_{k=0}^{\infty}\frac{1}{27^k}\left(\frac{3}{3k+1}+\frac{1}{3k+2}\right)$$

$$\frac{25}{2}\log\left(\frac{781}{256}\left(\frac{57-5\sqrt{5}}{57+5\sqrt{5}}\right)^{\sqrt{5}}\right) = \sum_{k=0}^{\infty}\frac{1}{5^{5k}}\left(\frac{5}{5k+2}+\frac{1}{5k+3}\right)$$

Is There a Base-10 Formula for Pi?

Note that there is both a base-2 and a base-3 BBP-type formula for π^2. Base-2 and base-3 formulas are also known for a handful of other constants.

Question: Is there any base-n BBP-type formula for π, where n is NOT a power of 2?

Answer: No. This is ruled out in a new paper by Jon Borwein, David Borwein and Will Galway.

This does not rule out some completely different scheme for finding individual non-binary digits of π.

Normal Numbers

- A number is b-normal (or “normal base b”) if every string of m digits in the base-b expansion appears with limiting frequency b^{-m}.
- Using measure theory, it is easy to show that almost all real numbers are b-normal, for any b.
- Widely believed to be b-normal, for any b:
  - π = 3.1415926535…
  - e = 2.7182818284…
  - Sqrt(2) = 1.4142135623…
  - Log(2) = 0.6931471805…
  - All irrational roots of polynomials with integer coefficients.

But to date there have been no proofs for any of these.
Proofs have been known only for contrived examples, such as C = 0.12345678910111213…

A Connection Between BBP Formulas and Normality

Consider the “chaotic’” sequence defined by $x_0 = 0$, and

$$x_n = \left\{2x_{n-1} + \frac{1}{n}\right\}$$

where { } denotes fractional part as before.

Result: log(2) is 2-normal if and only if this sequence is equidistributed in the unit interval.

In a similar vein, consider the sequence $x_0 = 0$, and

$$x_n = \left\{16x_{n-1} + \frac{120n^2 - 89n + 16}{512n^4 - 1024n^3 + 712n^2 - 206n + 21}\right\}$$

Result: π is 16-normal if and only if this sequence is equidistributed in the unit interval.

A Class of Provably Normal Constants

Crandall and I have also shown (unconditionally) that an infinite class of mathematical constants is normal, including

$$\begin{aligned}\alpha_{2,3} &= \sum_{k=1}^{\infty} \frac{1}{3^k 2^{3^k}} \\ &= 0.041883680831502985071252898624571682426096\ldots_{10} \\ &= \mathtt{0.0AB8E38F684BDA12F684BF35BA781948B0FCD6E9E0}\ldots_{16}\end{aligned}$$

$\alpha_{2,3}$ was proven 2-normal by Stoneham in 1971, but we have extended this to the case where (2,3) are any pair (p,q) of relatively prime integers. We also extended to uncountably infinite class, as follows [here r_k is the k-th bit of r in (0,1)]:

$$\alpha_{2,3}(r) = \sum_{k=1}^{\infty} \frac{1}{3^k 2^{3^k + r_k}}$$

A "Hot Spot" Lemma for Proving Normality

Recently Michal Misiurewicz and DHB were able to prove normality for these alpha constants very simply, by means of a new result that we call the "hot spot" lemma, proven using ergodic theory techniques:

Hot Spot Lemma: Let {} denote fractional part. Then x is b-normal if and only if there is no y in [0,1) such that

$$\liminf_{m\to\infty} \limsup_{n\to\infty} \frac{\#_{0\le j<n}(|\{b^j x\} - y| < b^{-m})}{2nb^{-m}} = \infty$$

Paraphrase: x is b-normal iff it has no base-b hot spots.

Hot Spot Examples

Consider 1/28. Successive shifts of the decimal expansion are:

0.142857142857142857142857142857142857 …

0.428571428571428571428571428571428571428 …

0.285714285714285714285714285714285714285714 …

0.857142857142857142857142857142857142857142857 …

...

Note that 1/7, 2/7, 3/7, 4/7, 5/7, 6/7 are "hot spots" – small intervals around these values are visited much too often by the sequence of shifted decimal expansions. Thus 1/28 is not a 10-normal number.

Similarly, consider the irrational constant

0.1001000010000001000000001000000000001… [1 at position n^2].

Note that successive shifts of the decimal expansion visit small neighborhoods of zero much too often, so this constant is not a 10-normal number (0.1, 0.01, 0.001, 0.0001, ... also are hot spots).

The BBP Sequence for $\alpha_{2,3}$

The BBP sequence for $\alpha_{2,3}$ can be seen to be

$0,$ repeated 3 times,

$\frac{1}{3}, \frac{2}{3},$ repeated 3 times,

$\frac{4}{9}, \frac{8}{9}, \frac{7}{9}, \frac{5}{9}, \frac{1}{9}, \frac{2}{9},$ repeated 3 times,

$\frac{13}{27}, \frac{26}{27}, \frac{25}{27}, \frac{23}{27}, \frac{19}{27}, \frac{11}{27}, \frac{22}{27}, \frac{17}{27}, \frac{7}{27}, \frac{14}{27}, \frac{1}{27}, \frac{2}{27}, \frac{4}{27}, \frac{8}{27}, \frac{16}{27}, \frac{5}{27}, \frac{10}{27}, \frac{20}{27},$

repeated 3 times, etc.

Note how this sequence very evenly fills the unit interval.

Proof that $\alpha_{2,3}$ is 2-Normal Using the Hot Spot Lemma

Suppose we are given some half-open interval [c, d), and let (x_j) be the BBP sequence for α. Observe that if $\{2^j \alpha\}$ is in [c,d), then x_j is in [c-1/(2j), d+1/(2j)). Let n be any integer > $1/(d-c)^2$, and let 3^p be such that $3^p <= n < 3^{p+1}$. Let m = int [1/(d-c) + 1]. Now note that for j >= m, we have [c - 1/(2j), d + 1/(2j)) is a subset of [3d/2 - c/2, 3c/2 -d/2). Since the length of this latter interval is 2(d-c), the number of multiples of $1/3^p$ that it contains is either int [2 x 3^p (d-c)] or one greater than this value. Thus there can be at most three times this many j's less than n for which x_j is in [3d/2-c/2, 3c/2-d/2). Therefore we can write

$$\begin{aligned}\frac{\#_{0\le j<n}(\{2^j\alpha\}\in[c,d))}{n(d-c)} &\le \frac{m+\#_{m\le j<n}(x_j\in[3d/2-c/2,3c/2-d/2))}{n(d-c)}\\ &\le \frac{m+3(2\cdot 3^p(d-c)+1)}{n(d-c)}<8.\end{aligned}$$

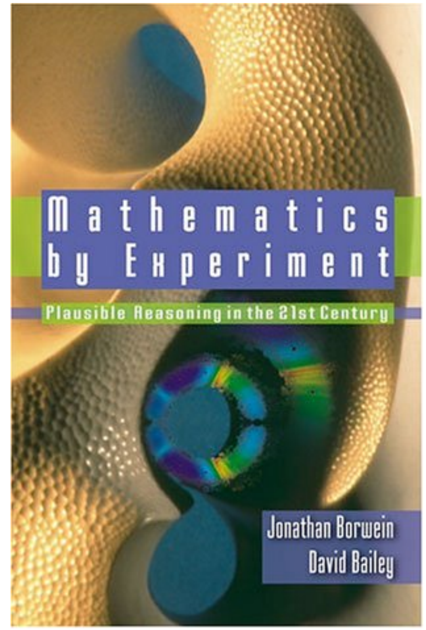
Thus by the hot-spot theorem, α is 2-normal.

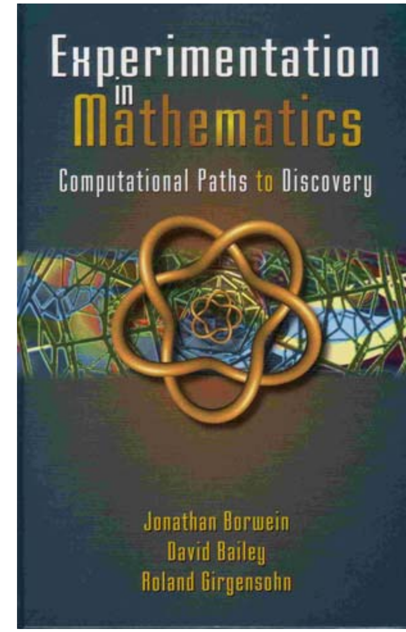
Two New Books on Experimental Mathematics

Vol. 1: Mathematics by Experiment: Plausible Reasoning in the 21st Century

Vol. 2: Experiments in Mathematics: Computational Paths to Discovery

Authors: Jonathan M Borwein and David H Bailey, with Roland Girgensohn for Vol. 2.





A "Reader's Digest" condensed version is available FREE at **http://www.experimentalmath.info**